PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, C12P 19/34, C07H 21/04

A2

(11) International Publication Number: WO 96/15264

(43) International Publication Date: 23 May 1996 (23.05.96)

(21) International Application Number: PCT/GB95/02654

(22) International Filing Date: 13 November 1995 (13.11.95)

(30) Priority Data:
9422891.3 12 November 1994 (12.11.94) GB

(71) Applicant *(for all designated States except US)*: JSD TECHNOLOGIES LTD. [GB/GB]; Premier House, 29 Rutland Street, Leicester LE1 1RE (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: STEWART, Gordon, Sydney, Anderson, Birnie [GB/GB]; 14 James Avenue, Loughborough LE11 0QL (GB). FRANCIS, Kevin, Patrick [GB/GB]; 40 Elizabeth Avenue, North Hykeham, Lincoln, Lincolnshire LN6 9RR (GB).

(74) Agent: PENNANT, Pyers; Stevens, Hewlett & Perkins, 1 Serjeants' Inn, Fleet Street, London EC4Y 1LL (GB).

(81) Designated States: US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: METHOD FOR IDENTIFYING NUCLEIC ACID SEQUENCES FROM DIFFERENT BIOLOGICAL SOURCES THROUGH AMPLIFICATION OF THE SAME

(57) Abstract

A method of identifying a nucleic acid sequence in a biological sample comprises using a pair of universal oligonucleotide primers to amplify the nucleic acid sequence and characterising the amplification reaction products. Preferred universal primers are derived from conserved regions of cold shock or Y-box proteins and hybridise to the genes that code for the peptide sequences GXVKWFNXXKGFGFI and GPXAXNVTXX.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

Method for identifying nucleic acid sequences from different biological sources through amplification of the same

The present invention relates to improvements in and relating to the amplification and identification of nucleotide base sequences, and in particular but not exclusively to improvements in and relating to the amplification and identification of nucleotide base sequences using the polymerase chain reaction (PCR).

The PCR is a useful technique for amplifying genetic sequences. One application of this is the amplification of target gene sequences in biological samples from for example environmental, food and medical sources etc. to allow identification of causative, pathogenic, spoilage or indicator organisms present in the sample.

Conventionally, the PCR is carried out using primers chosen or produced to amplify a particular target gene sequence within a given organism. Consequently, currently the PCR can only be used to amplify a single predetermined gene sequence at any one time and hence to confirm the presence or absence of that target sequence and the corresponding organism. Thus, when the identity of one or more organisms in a sample is to be established, it is necessary to guess or assume what sequences/organisms may be present, identify and obtain specific primers that are operable to amplify that sequence or a sequence indicative of that organism, and then conduct the experiment. Such procedures are time consuming and expensive, and somewhat unreliable.

It is an object of the present invention to obviate or mitigate one or more of the above disadvantag s.

Throughout this specification the terms "unidentified sequence" and "unidentified nucleotide sequence" are used to mean that the sequence has not yet been identified in the particular sample being analysed, not necessarily one which is wholly unknown. It may be that once an unidentified sequence has been amplified and characterised it is one with which scientists are familiar.

In one aspect the invention provides a method of identifying a nucleic acid sequence in a biological sample, which method comprises using at least one pair of oligonucleotide primers to amplify the nucleic acid sequence, and characterising the amplification reaction products, characterised in that the or each pair of oligonucleotide primers is operable to enable the amplification of nucleic acid sequences from different biological sources.

The amplification method is not material to the invention. Preferably amplification is effected by the polymerase chain reaction (PCR).

The method by which amplification reaction products are characterised is also not material to the invention. Such characterisation may be effected by known techniques, e.g. sequencing, or by the use of single-strand conformation polymorphism (SSCP) analysis, or by the use of a labelled probe which binds to the amplification reaction products. One suitable system is that marketed by Perkin Elmer under the trademark TaqMan which is described in more detail below.

The method of the invention is characterised by the use of at least one pair of oligonucleotide primers operable to enable the amplification of nucleic acid sequences from different biological sources. These so-called universal primers can be selected from regions of genes which are largely or completely conserved in many different biological organisms. Pairs of univ rsal primers may be found by routine search in various different genes. To the best of the inv ntors' knowledge, the use

of such pairs of universal primers in order to identify the biological source of a particular nucleic acid sequence has not previously been proposed.

In another aspect the invention provides a kit for identifying a nucleic acid sequence in a biological sample by the method described which kit comprises reagents, including at least one pair of oligonucleotide primers operable to enable the amplification of nucleic acid sequences from different biological sources, for amplifying the nucleic acid sequence, and means for characterising the amplification reaction products.

Studies have revealed that the proteins involved in the cold-shock response in prokaryotic organisms are remarkably conserved throughout different organisms. Such conservation also extends to related proteins in eukaryotic organisms, for example the Y-box proteins. Very highly conserved regions of the amino acid sequences of these proteins have been identified, and recognised by the inventors as potential sites for universal amplification primers.

Preferably a pair of oligonucleotide primers includes a first oligonucleotide primer which is substantially complementary to a nucleic acid sequence that codes for the peptide sequence GXVKWFNXXKGFGFI (SEQ ID NO: 1) where X is any amino acid, or for a fragment thereof that includes at least four identified amino acids.

Herein, peptide sequences are shown in the N to C terminus direction. Oligonucleotides are shown in the 5' to 3' direction. Reference to an amino acid in a peptide sequence is to be understood as a reference to the residue of that amino acid. Reference to a nucleotide in an oligonucleotide sequence is to be understood as a reference to a residue of that nucleotide. An identified amino acid is one not indicated by the symbol X.

Preferably the said first oligonucleotide primer has the s quence GGTANAGTAAAATGGTTNAACNC (SEQ ID NO: 2) where N is any nucleotide, or a fragment thereof containing at least twelv nucleotides.

Preferably a second oligonucleotide primer is substantially complementary to a nucleic acid sequence that codes for the peptide sequence GPXAXNVTXX (SEQ ID NO: 3) where X is any amino acid, or for a fragment thereof that includes at least four identified amino acids. Preferably the said second oligonucleotide primer has the sequence GGTTACGTTANCNGCTNNNGGNCC (SEQ ID NO: 4) where N is any nucleotide, or a fragment thereof containing at least twelve nucleotides.

Preferably a third oligonucleotide primer used is substantially complementary to a nucleic acid sequence that codes for the peptide sequence DVFVHFSAIQ (SEQ ID NO: 5).

Preferably the said third oligonucleotide primer used has the sequence GATGTATTCGTACATTTCTCTGCTATCC (SEQ ID NO 6) or its reverse complement GGATAGCAGAGAAATGTACGAATACATC (SEQ ID NO: 7) or a fragment of either containing at least twelve nucleotides or an oligonucleotide having at least 80% homology therewith. The third oligonucleotide primer or its reverse complement forms a pair with either the second or the first oligonucleotide primers, respectively. Alternatively, the first oligonucleotide primer forms a pair with the second oligonucleotide primer.

In another aspect the invention provides as new products, a pair of oligonucleotide primers comprising the first and one of the other two oligonucleotides, and a pair of oligonucleotide primers comprising the second and third oligonucleotides, and a set of primers comprising all three oligonucleotides.

Reference is directed to the accompanying drawings in which Figure 1 is a comparison of the aligned amino acid sequences of the major cold shock proteins of various bacteria and of a number of eukaryotic Y-box proteins.

Figure 2 shows the sequences of three conserved regions of genes which code for these proteins, and includes three oligonucleotides

whose sequences or reverse complementary sequences are suitable for use as primers according to this invention.

Figure 3 is a schematic representation of an oligonucleotide primer set in use in the amplification of a nucleic acid sequence according to this invention.

Figures 4 and 5 show the front and back oligonucleotide primers from Figure 2, and how they relate to conserved peptide sequences of cold shock and Y-box proteins.

Figure 1 shows alignments of amino acid sequences of published cold shock and Y-box proteins. (: = space in sequence; - = conserved amino acid). E.c = *Escherichia coli*, B.s = *Bacillus subtilis*, B.g = *Bacillus globigii*, B.gl = *Bacillus globisporus*, B.c = *Bacillus caldolyticus*, B.st = *Bacillus stearothermophilus*, S.c = *Streptomyces calvuligers*, H.s = *Homo sapiens*, R.n = *Rattus norvegicus*, X.l = *Xenopus laevis*, A.t = *Arabidopsis thaliana*. The data clearly demonstrates that these proteins share large regions of homology. By identifying the most highly conserved regions within these alignments, oligonucleotide primers were designed which, when used in a PCR reaction using chromosomal DNA from various bacteria, facilitate the amplification of short DNA sequences (~180 bp) possessing regions of homology with published major cold shock protein sequences.

Figure 2 focuses on three highly conserved regions of Figure 1, and shows how the genes which code for these proteins are also highly conserved, both across various bacterial species and also in the frog. These conserved regions have been used to design three degenerate primers;

- a first or front primer which hybridises to the active coding strand of the DNA;

- a middl or third oligonucleotid which hybridis s to the coding or the non-coding strand of the DNA; and

- a back or second primer which hybridis s to th non-coding strand of the DNA.

In these degenerate oligonucleotides, a capital letter designates a nucleotide that is conserved throughout; a small letter designates a nucleotide that is not universally conserved; and a dot indicates that the nucleotide shown below it can be changed at will.

The three oligonucleotides shown in Figure 3 are indeed universal primers. Using them, the inventors have successfully amplified DNA sequences in biological samples derived from bacteria, eukaryotes including rat, frog and man, and plant tissue. For example, nucleic acid sequences have been amplified from the following micro-organisms:-

Bacillus subtilis
Xanthomonas campestris
Salmonella typhimurium
Micrococus luteus
Chromobacterium violaceum
Agrobacterium lumefaciens
Azospirillus
Rahnella aquaticus
Pseudomonas putida
Serratia liquifaciens
Serratia marcescens
Pseudomonas fluorescens
Staphylococcus aureus
Pseudomonas aeruginosa
Pseudomonas mesoacidophilus
Photobacterium leiognathi
Vibrio fischeri
Xenorhabdus luminescens
Listeria monocytogenes

Bacillus stearothermophilus

Thermophilus aquaticus

Yersinia enterocolitica

Amplification was performed under conventional PCR conditions, using a touchdown technique to minimise interference from impurities. The following temperature profile was used:-

a) 94°C for 6 minutes.

b) 94°C for 30 seconds.

60 - 51°C for 30 seconds.

72°C for 30 seconds.

b) repeated 20 times with the annealing temperature reduced by one degree every second cycle.

c) 94°C for 30 seconds

50°C for 30 seconds

72°C for 30 seconds

c) cycle repeated 20 times.

Figure 3 shows front and middle primers of Figure 2 hybridised to the coding strand of a nucleic acid, and the reverse complement back primer from Figure 2 hybridised to the non-coding strand of the DNA. The front and back primers form a pair of oligonucleotide primers according to the invention. The middle and back primers form another pair of oligonucleotide primers. The three primers together form a set according to the invention.

Figure 4 shows the highly conserved amino acid sequence of the cold shock proteins used to produce the front primer of the primer pair. Figure 5 shows the highly conserved peptide sequence of the cold shock proteins used to produce the back primer of the primer pair. Each X in the sequence indicates a position in the sequence at which the precise nature of the amino acid is not particularly important sinc these positions are less conserved than the others. Consequently th nature of the primer at the

corresponding position can vary without d triment to the efficacy of the respective primer.

Figure 4 also shows the front oligonucleotide primer of the universal primer pair, alongside the particularly conserved peptide sequence from which it is derived. It should be noted that the conservation of the peptide sequence is greater than that of the corresponding nucleotide sequence due to "wobble", i.e. amino acids can be coded for by more than one codon triplet. Thus, the primer is one of a number of possible primers that are hybridisable to the DNA sequence actually coding for the conserved sequence GXVKWFNXXKGFGFI (SEQ ID NO: 1). This primer is complementary to and hence operable to hybridise to a first primer binding region on the coding or sense strand of a cold shock protein gene sequence, as shown in Figure 3.

Figure 5 shows the back oligonucleotide primer of the pair, alongside the highly conserved peptide sequence from which it is derived. The back primer is operable to hybridise with a second primer binding region on the noncoding or anti-sense strand of a cold shock protein gene sequence, again as shown in Figure 3.

The primer binding regions are separated by a generally less conserved series of about 180 base pairs. This less well conserved part of the gene encodes for a cold-shock response protein, or a related or similar protein such as a Y-box protein (eukaryotes). The primer binding regions are sufficiently conserved to enable the universal primers to recognise and hybridise to the DNA of a large number, possibly all or almost all, organisms, including both prokaryotic and eukaryotic organisms and thereby enable amplification of the intervening sequences from these organisms as will be described. The primer pair is at least expected to operate with very large classes of organisms and thus to be "universal" in this sense. The gene howev r does have a degree of peculiarity in ach differ nt type of organism, such that characterisation of the amplification

products made with the primers and identification of the sequence, the RNA, protein and source organism can be made (see later).

In use, the primers are operable in a generally conventional polymerase chain reaction to amplify a gene encoding for a cold shock or related protein or indeed any sequence bounded by the highly conserved primer binding regions. It is preferable that the reaction is conducted using a substantially DNA free thermophilic transcription enzyme to mitigate "background" amplification i.e. amplification of exogenous DNA associated with the enzyme. However, use of an improved polymerase chain reaction method in accordance with the present invention may render this requirement less crucial (see below). A suitable enzyme is replitherm.

One particular application of the universal primer pair is the identification of unidentified source organisms in a biological sample. For example, food, environmental and medical samples can be analysed to identify the nature of a particular causative, pathogenic spoilage and/or indicator organism.

Conventionally sequence specific and hence organism-specific primers have had to be chosen or produced to enable such identification procedures. One significant disadvantage of this, which the present invention obviates, is the need for the investigator to guess or assume what organism may be involved, and test for it. If that organism is not involved, then the investigator will have to guess again. This has to be repeated until the investigator has correctly guessed or assumed the identity of the organism involved and the test has proved this to be the case. It is feasible that the organism involved is not identified as a result of such laborious techniques. Moreover, it may be that more than one organism is involved, a fact that may be missed using conventional techniques, but which should not be missed using the present invention.

The present invention allows an investigator simply to conduct the polymerase chain reaction using the universal primers to

amplify all sequences present in the sample bound by the relevant highly conserved primer binding regions. The polymerase reaction products (the amplified sequences) can then be characterised, for instance by sequencing them or using high resolution gel electrophoresis, to enable identification of the sequence amplified. It is preferable that techniques ar adopted which obviate the need to sequence the amplification products, since sequencing is a relatively laborious process.

Therefore, in accordance with the present invention the gel pattern produced using electrophoretic techniques to separate the reaction products produced using the primers may be compared against a data bank of gel patterns produced from known sequences and organisms, such that identification of the sequence and hence source organism in the sample being analysed can be readily ascertained by matching it with one in the data bank. If there is no corresponding pattern on the data bank then the products may have to be sequenced, or identified by another appropriate method. It will be appreciated that determination of the identity of the sequence and source organism by such comparative procedures is highly efficient, particularly when the comparison is conducted by a computer.

The use of fluorescent labels to label the PCR products facilitates the identification procedure.

The front and back and middle primers may be labelled with fluorescent labels (Figure 3; e.g. FAM, TAMRA, JOE as shown in Figure 2). Thus, when the reaction products are separated on a gel, the positions of the reaction products can be established, preferably by using apparatus operable to detect the labels. Such apparatus may employ laser technology to detect the labels in a conventional manner.

Once the position of the reaction products have been established, they can be calibrated or standardised to take into account varying parameters such as the type of label used, temperature etc., and

then compared, by computer with known patterns. The differences in the less conserved intervening sequences will cause the reaction products of different sequences to move through the separating gel at different rates, as a result of differences in size, charge, etc. It is therefore possible to distinguish between different sequences, to an accuracy of one base difference, as in the fluorescence-based PCR single strand conformation polymorphism analysis (PCR-SSCP) [Ellison *et al*, BioTechniques, Vol 15 No 4 (1993) p684; and Ellison *et al*, BioTechniques, Vol 17 No 4, 1994, p 742], and hence to distinguish and identify the sequence and source organism in this way.

If the procedure results in a pattern which does not compare with a known pattern, the investigator is still alerted to the presence of an organism, which can then be further investigated by other procedures.

A preferred method of characterising the amplification reaction products involves use of the system marketed by Perkin Elmer under the trademark TaqMan. This is a PCR assay system that uses the hydrolysis of a fluorogenic probe to monitor the extent of amplification. The probe consists of an oligonucleotide labelled with both a fluorescent reporter and a quencher dye. During the PCR process, this probe is cleaved by the 5' nuclease activity of taq DNA polymerase if and only if it hybridises to the segment being amplified. Cleavage of the probe generates an increase in the fluorescence intensity of the reporter dye. Thus, amplification of a specific product can be detected by simply measuring fluorescence after PCR. A separate probe is required for each bacterial or other species to be investigated. But once synthesised, the probes are very quick and easy to use.

As previously mentioned, an improved method of amplifying and identifying a nucleotide sequence comprises using a primer set of at least three cooperable primers to amplify the sequence and at least one portion thereof. The primers of the primer s t are preferably the thre

universal primers as described above. However, it is to be appreciated that the improved method is operable with non-universal primers, for example to increase the fidelity of the PCR-SSCP analysis mentioned above.

The improved method likewise increases the specificity and fidelity of the characterisation and identification methods using the universal primer set described above.

Generally conventional PCR methodology is employed to simultaneously amplify different parts of the sequence using these primers. Two principal reaction products are produced, one between the front and back primers, and one between either the middle and back primers or the front and middle primers depending on which of the middle primers is used.

Thus two different distinguishable sets of reaction products are produced for each sequence amplified. Therefore, the gel pattern produced by separation of these is likely to be more specific and therefore more distinguishable for a given sequence and hence organism than a gel pattern of reaction products produced using only one primer pair.

Each of the reaction products can have a different label to enable them to be identified on the gel.

It is to be appreciated that use of three primers, or more would be in accordance with the present invention. Indeed the more primers that were used, the greater the fidelity and specificity of the technique. The use of four primers is envisaged, where front and back primers plus two middle primers as described above are used. In this case, the two middle primers will be of opposite sense but each will correspond to a different part of the conserved region to avoid formation of primer dimers.

Using the present invention, the investigator can readily resolve the general question, "What is present?", rather than ask the specific question "Is organism X present?".

The invention therefore provides a considerably simpler and more efficient means to identify nucleotide sequences and source organisms in biological samples.

It is to be appreciated that proteins other than the cold shock proteins (and related proteins) may comprise suitable highly conserved regions appropriate for producing universal primers, and that the present invention extends to primer pairs, sets and methods suitable for use in relation thereto. For example, ribosomal proteins, sigma subunits and RsmA comprise regions that appear highly conserved throughout a range of sources, and therefore would appear to provide appropriate sites for suitable universal primers.

Moreover, universal primers may be used in reactions other than PCR.

CLAIMS

1. A method of identifying a nucleic acid sequence in a biological sample, which method comprises using at least one pair of oligonucleotides primers to amplify the nucleic acid sequence, and characterising the amplification reaction products, characterised in that the or each pair of oligonucleotide primers is operable to enable the amplification of nucleic acid sequences from different biological sources.

2. A method as claimed in claim 1, wherein the amplification is performed by the polymerase chain reaction.

3. A method as claimed in claim 1 or claim 2, wherein the characterising is performed by sequencing or by single-strand conformation polymorphism analysis or by the use of a labelled probe which binds to the amplification reaction products.

4. A method as claimed in any one of claims 1 to 3, wherein th oligonucleotide primers used hybridise to a gene which codes for a cold shock protein or a eukaryotic Y-box protein.

5. A method as claimed in any one claims 1 to 4, wherein a first oligonucleotide primer used is substantially complementary to a nucleic acid sequence that codes for the peptide sequence GXVKWFNXXKGFGFI (SEQ ID NO: 1) where X is any amino acid, or for a fragment thereof that includes at least four identified amino acids.

6. A method as claimed in claim 5, wherein another oligonucleotide primer used is substantially complementary to a nucleic acid s quence that codes for the p ptide sequ nce GPXAXNVTXX (SEQ ID NO: 3) wher

X is any amino acid, or for a fragment th reof that includes at l ast four identified amino acids.

7. A method as claimed in claim 5, wherein the oligonucleotide primer used has the sequence GGTANAGTAAAATGGTTNAACNC (SEQ ID NO: 2) where N is any nucleotide, or a fragment thereof containing at least twelve nucleotides.

8. A method as claimed in claim 6, wherein the oligonucleotide primer used has the sequence GGTTACGTTANCNGCTNNNGGNCC (SEQ ID NO: 4) where N is any nucleotide, or a fragment thereof containing at least twelve nucleotides.

9. A method as claimed in any one of claims 5 to 8, wherein another oligonucleotide primer used has the sequence GATGTATTCGTACATTTCTCTGCTATCC (SEQ ID NO: 6) or its reverse complement GGATAGCAGAGAAATGTACGAATACATC (SEQ ID NO: 7) or a fragment of either containing at least twelve nucleotides

10. A pair of oligonucleotide primers, operable to enable the amplification of nucleic acid sequences from different biological sources, comprising a first oligonucleotide primer substantially complementary to a nucleic acid sequence that codes for the peptide sequence GXVKWFNXXKGFGFI (SEQ ID NO: 1) where X is any amino acid, or for a fragment thereof that includes at least four identified amino acids, and another oligonucleotide primer substantially complementary to a nucleic acid sequence that codes for the peptide sequence GPXAXNVTXX (SEQ ID NO: 3)where X is any amino acid, or for a fragment thereof that includes at least four identified amino acids.

11. A set of oligonucleotide primers, operable to enable the amplification of nucleic acid sequences from different biological sources, comprising a first oligonucleotide primer having the sequence GGTANAGTAAAATGGTTNAACNC (SEQ ID NO: 2) where N is any nucl otid , or a fragment thereof containing at least twelve nucl otides, and at least one

other oligonucleotid primer having a sequence selected from (i) GGTTACGTTANCNGCTNNNGGNCC (SEQ ID NO: 4) where N is any nucleotide, or a fragment thereof containing at least twelve nucleotides and (ii) GATGTATTCGTACATTTCTCTGCTATCC (SEQ ID NO: 6) or its reverse complement GGATAGCAGAGAAATGTACGAATACATC (SEQ ID NO: 7) or a fragment of either containing at least twelve nucleotides

12. A kit for identifying a nucleic acid sequence in a biological sample by the method of any one of claims 1 to 9, which kit comprises reagents, including at least one pair of oligonucleotide primers operable to enable the amplification of nucleic acid sequences from different biological sources, for amplifying the nucleic acid sequence, and means for characterising the amplification reaction products.

13. A kit as claimed in claim 12, wherein the at least one pair of oligonucleotide primers is a pair or set of oligonucleotide primers as claimed in claim 10 or claim 11.

Fig.1.

```
E.c  CspA      MSGKMTGIVKWFNADKGFGFITPDDGSKDVFVHFSAIQN
E.c  CspB      --N----L----------S-V---------------
E.c  CspC      -:A-IK-Q-----ES--------A----------G
E.c  CspD      -:::EK-T-----NA------C-EG-GE-I-A-Y-T--M
B.s  CspB      -:::LE-K-----SE------EV:E-QD----------G
B.g  CspBgl          -----SE--L---EV:E-QD----------G
B.gl CspBgs          I----SE------EV:E-QD----------G
B.c  CspBca    -:::QR-K-----NE--Y---EVEG--:------T---G
B.st CspBst    -:::QR-K-----NE--Y---EVEG--:------T---G
S.c  SC7.0     -:::A--T------E------AQ-G-GP-----Y---NA
H.s  YB-1      IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KK
R.n  EFI       IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KK
X.l  FRGY-1    IAT-VL-T-----VRN-Y---NRN-TKE-----QT--KK
A.t  GRP-2     GGERRK-S----DTQ----------GD-L---Q-S-RS

E.c  CspA      ::::DGYKSLDEGQKVSFTI:ESGAKGPAAGNVTSL
E.c  CspB      ::::-N-RT-F-----T-S-:---------A--IITD
E.c  CspC      ::::N-F-T-A---N-E-E-:QD-Q-----V---AI
E.c  CspD      :::::---RT-KA--S-Q-DV:EQ-P--NE-SVIVPVEVEVAAVA
B.s  CspB      ::::E-F-T-E---A---E-:VE-NR--Q-A---KEA
B.g  CspBgl    ::::E-F-T-ES--A---E-:VE-NR--Q-A
B.gl CspBgs    ::::E-FCT-E---A---E-:VE-NR--Q-A
B.c  CspBca    ::::E-F-T-E--EA---EQ:VE-NR--Q-AV-L
B.st CspBst    ::::E-FST-E--EA---EQ:VE-NR--Q-AV-L
S.c  SC7.0     ::::T-FR--E-N-V-N-DV:TE-E:--Q-E--SPA
H.s  YB-1      NNPRKYLR-VGD-ET-E-DV:VE-E--EE-A---GPG
R.n  EFI       NNPRKYLR-VGD-ET-E-DV:VE-E--AE-A---GPG
X.l  FRGY-1    NNPRKYLR-VGD-ET-E-DV:VE-E--AE-A---GPG
A.t  GRP-2     ::::E-FR--AAEE--E-EVEIDN-NR-K-ID-SGPD
```

Fig.2.

FRONT	*E. c cspA*	GGTATCGTAAAATGGTTCAACGC	20/23	(SEQ ID NO: 8)
	E. c cspB	GGTTTAGTAAAATGGTTTAACGC	21/23	(SEQ ID NO: 9)
	E. c cspC	GGTCAGGTTAAGTGGTTCAACGA	17/23	(SEQ ID NO: 10)
	S. t cspS	GGTATCGTAAAATGGTTCAACGC	20/23	(SEQ ID NO: 11)
	S. t cspB	GGTTTAGTAAAATGGTTTAACCC	20/23	(SEQ ID NO: 12)
	B. s cspB	GGTAAAGTAAAATGGTTCAACTC	21/23	(SEQ ID NO: 13)
	L. m cspL	GGTACAGTAAAATGGTTTAACGC	22/23	(SEQ ID NO: 14)
	P. m cspP	GGAAAAGTAAAATGGTTTAACGC	22/23	(SEQ ID NO: 15)
	PRG Y1	GGGACAGTCAAATGGTTTAATGT	18/23	(SEQ ID NO: 16)
	UNI	GGtaaaGTaAAaTGGTTtAAcgc	FAM	(SEQ ID NO: 17)
MIDDLE	*E. c cspA*	GATGTATTCGTTCATTTCTCTGCTATTC	26/28	(SEQ ID NO: 18)
	E. c cspB	GATGTGTTTGTGCATTTTTCTGCGATTC	22/28	(SEQ ID NO: 19)
	E. c cspC	GATGTGTTCGTACACTTCTCCGCTATCC	25/28	(SEQ ID NO: 20)
	S. t cspS	GATGTGTTCGTACACTTCTCTGCTATTC	25/28	(SEQ ID NO: 21)
	S. t cspB	GATGTGTTCGTACACTTCTCTGCTATCC	26/28	(SEQ ID NO: 22)
	B. s cspB	GATGTATTCGTTCATTTCTCTGCTATTC	26/28	(SEQ ID NO: 23)
	L. m cspL	GATGTATTCGTACATTTCAGCGCTATCC	25/28	(SEQ ID NO: 24)
	P. m cspP	GATGTATTCGTTCACTTCTCAGCTACTC	23/28	(SEQ ID NO: 25)
	PRG Y1	GATGTGTTTGTACACCAAACTGCCATCA	19/28	(SEQ ID NO: 26)
	UNI	GATGTaTTcGTaCAtttctctGCtAtcc	TAMRA	(SEQ ID NO: 27)
BACK	*E. c cspA*	GGCCCGGCAGCTGGTAACGTAACC	19/24	(SEQ ID NO: 28)
	E. c cspB	GGTCCTGCAGCAGCAAATGTCATC	18/24	(SEQ ID NO: 29)
	E. c cspC	GGTCCGGCAGCTGTTAACGTAACA	18/24	(SEQ ID NO: 30)
	S. t cspS	GGCCCGGCAGCTGGTAACGTAACC	19/24	(SEQ ID NO: 31)
	B. s cspB	GGACCACAAGCTGCTAACGTTACT	19/24	(SEQ ID NO: 32)
	L. m cspL	GGACCTCAAGCAGCTAACGTTCAA	19/24	(SEQ ID NO: 33)
	PRG Y1	GGTGCAGAGGCAGCTAATGTAACT	18/24	(SEQ ID NO: 34)
	UNI	GGacCtgaaGCaGctAAcGTAacc	JOE	(SEQ ID NO: 35)

Fig.3.

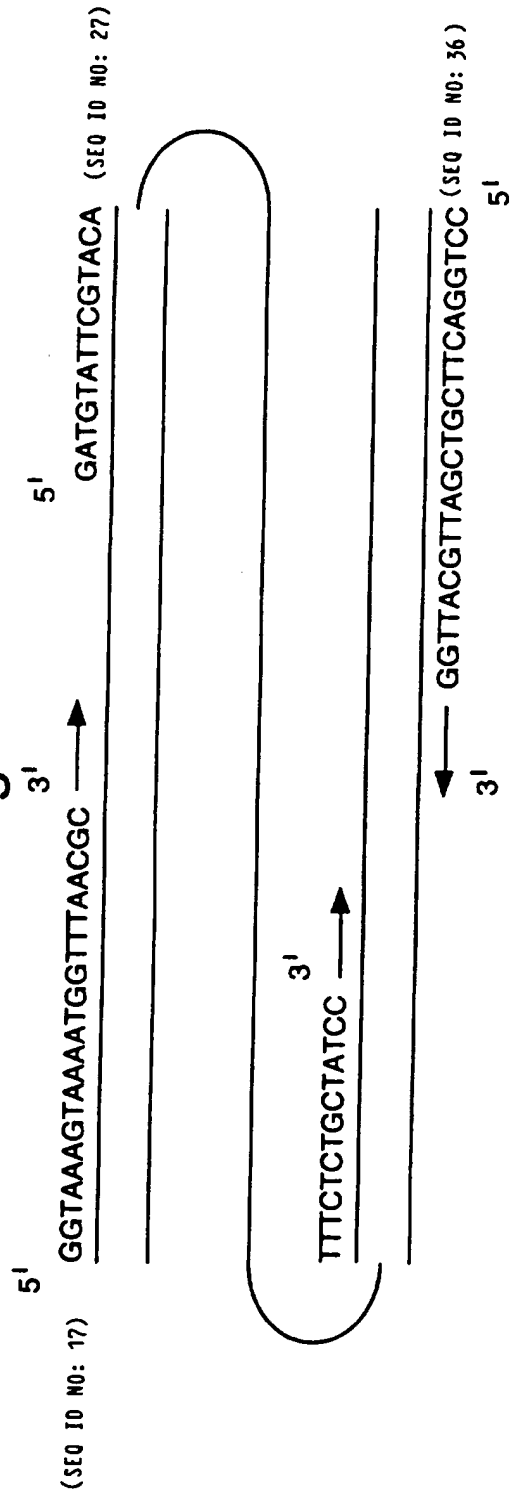


Fig.4.

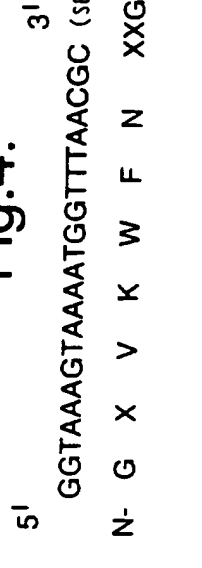
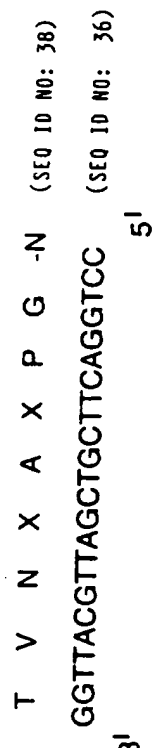


Fig.5.



RECTIFIED SHEET (RULE 91)
ISA/EP

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/GB 95/02654

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68 C12P19/34 C07H21/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,92 05277 (DAVIDSON AND BARTLETT) 2 April 1992 see page 28, line 17 - page 30, line 10 see page 33, line 6 - line 27; claims 1-22	1-3,12, 13
A	FR,A,2 636 075 (SOCIETE EUROPEENE DE BIOTECHNOLOGIE) 9 March 1990 see the whole document	1
A	EP,A,0 245 129 (INSTITUT PASTEUR) 11 November 1987 see claims 1-10	1
A	EP,A,0 438 115 (THE PERKIN-ELMER CORP.) 24 July 1991 see the whole document	1

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
22 April 1996	02. 05. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Osborne, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/GB 95/02654

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO,A,93 20237 (CAMBRIDGE NEUROSCIENCE) 14 October 1993	

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 95/02654

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9205277	02-04-92	CA-A-	2026264	27-03-92
		AT-T-	119943	15-04-95
		AU-B-	8515391	15-04-92
		DE-D-	69108249	20-04-95
		DE-T-	69108249	09-11-95
		EP-A-	0550491	14-07-93
		ES-T-	2072618	16-07-95
		FI-A-	931318	24-05-93
FR-A-2636075	09-03-90	NONE		
EP-A-245129	11-11-87	FR-A-	2596774	09-10-87
		AU-B-	612571	18-07-91
		AU-B-	7232587	20-10-87
		DE-A-	3776669	26-03-92
		WO-A-	8705907	08-10-87
		JP-T-	1500001	12-01-89
		KR-B-	9513772	15-11-95
		NO-B-	174780	28-03-94
		US-A-	5407797	18-04-95
EP-A-438115	24-07-91	AU-B-	651077	14-07-94
		AU-B-	6947591	25-07-91
		CA-A-	2033718	20-07-91
		US-A-	5298392	29-03-94
WO-A-9320237	14-10-93	AU-B-	4100793	08-11-93